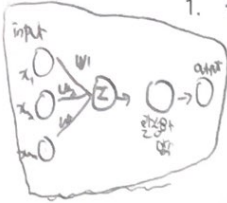


## 객체 Quiz 2020/10/29

학번: 201112159

성명: 우승제

### 1. 신경망이 훈련되는 원리를 설명하라



인공신경망은 두뇌의 신경세포를 모방하여 수학적으로 모델링한 것이며, 데이터를 학습시켜 분류 및 예측의 처리를 진행하는 것이다.

신경망이 훈련되는 원리는, 그림과 같이 새로운 입력데이터 각각의 입력값  $x$ 에 값을 가진 노드들이 임의의값을 갖는 가중치  $w$ 와 결합하여 새로운 노드를 생성하게 되는 원리로 학습된다.

### 2. 훈련데이터와 평가데이터란?

훈련데이터는 신경망 훈련시에 사용하고 평가데이터는 학습결과를 검증하는데 쓴다.

보통, 훈련데이터 샘플 수는 평가데이터 샘플 수보다 많게 한다.

### 3. 신경망의 분류와 회귀란?

분류는 이진분류와 다중분류로 나뉘며, 분류에서는 주로 단일층신경망과 이진분류는 2가지로 나눠 분류하는 것이다. 다중신경망 알고리즘을 쓴다.

(a) 종양의 유무, 성별

다중분류는 여러 가지 분류하는 것.

(b) 키, 몸무게에 Object를

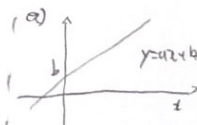
회귀는 연속하는 수치를 예측하는 것이다.

(c) 기온, 강수량

회귀에서는 주로 선형회귀 알고리즘을 쓴다.

1. 선형회귀는 선형모델로 수치를 예측한다.

2. 손실함수는 저등회귀함수가 없다.



### 4. 신경망의 과대적합을 설명하고, 완화시킬 수 있는 방법들을 기술하라.

과대적합은 Overfitting으로 모델이 학습데이터를 완벽히 학습하고, 과하게 맞추게

훈련데이터에 포함된 노이즈까지 학습한 상태를 말한다.

이를 완화시키는 방법으로는

1. 데이터양을 늘리기

2. 모델의 복잡도 줄이기

3. Regularization 적용하기

4. 드롭아웃 등에 신경망분류에 적용해준다.

등의 방법이 있다.

5. 단어 임베딩이란?

단어 임베딩이란, word를 R차원의 vector로 매핑해주는 것을 말한다.

단어  $\Rightarrow$  dense vector

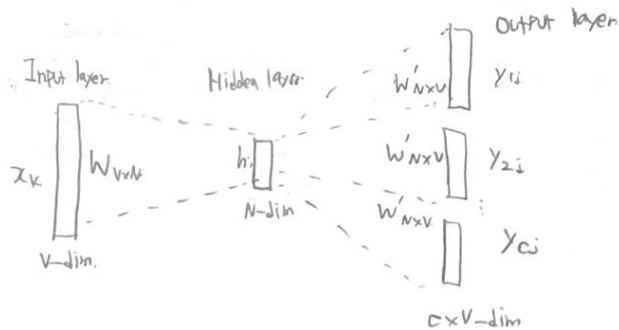
단어는 적절한 차원수에  $\rightarrow$  모든 것을 표현하기 위해 단어 임베딩을 한다.

ex) SVD.

6. word2vec 방법에 대해서 설명하라.

Word2vec은 단어 임베딩 중 한 가지 방법이다.

N/L을 계층마다의 학습속도와 성능을 비교하여 골라 붙여 준다.



word2vec은 은닉층이 하나인 뉴럴네트워크 구조이다.

여기서  $W, W'$ 은 word2vec의 학습값의 행렬이다.

입력층 - 은닉층, 은닉층 - 출력층을 잇는 가중치행렬이 모양이 transpose한 것을 알 수 있다.

중심단어와 주변단어는 '맞คู่'거나 '잘못'된 쌍으로 중심 단어를 '찾'거나 'w, w'를 텐서로 나타내어 인출시켜 준다.

7. Small IMDB dataset 을 사용해서 sentiment analysis 를 할 때, 기본적인 모델은 다음과 같다. (vectorize\_layer, vocab\_size, embedding\_dim 는 제공된다고 가정)
- 모델의 정확도를 향상 시키기 위해서 모델을 변경하라. (오직 모델만 변경 가능)

```
model = tf.keras.Sequential([
    vectorize_layer,
    layers.Embedding(vocab_size, embedding_dim),
    layers.GlobalAveragePooling1D(),
    layers.Dense(1)])
```

```
model = tf.keras.Sequential ([
```

```
vectorize_layer,
```

```
layers.Embedding(vocab_size, embedding_dim),
```

```
layers.GlobalAveragePooling1D(),
```

```
layers.Dense(1, activation='sigmoid')])
```

```
es = EarlyStopping (monitor='val_loss', min_delta=1e-4, verbose=1, patience=4)
```